

A Project Report On

Chemicals in Makeup

Submitted in partial fulfillment of the
requirement for the award of the degree

Bachelor of Computer Application (BCA)

Academic Year 2025 – 26

Soumya Bhatt
92320527006

Rabab Travadi
92300527070

Venisha Vora
92300527042

Internal Guide
Vinod KumarPal



Marwadi
University
Marwadi Chandarana Group



Faculty of Computer Applications (FCA)

Certificate

This is to certify that the project work entitled
Chemicals in Makeup
submitted in partial fulfillment of the requirement for
the award of the degree of
Bachelor of Computer Application
of the
Marwadi University
is a result of the bonafide work carried out by
Soumya Bhatt (92320527006)
Rabab Travadi (92300527070)
Venisha Vora (92300527042)

during the academic year 2025-26

Prof. Vinod Kumar Pal
Faculty Guide

Prof. Sunil Bajaja
HOD

Dr. Prof. R. Sridaran
Dean

DECLARATION

We hereby declare that this project work entitled **Chemicals in Makeup** is a record done by me.

We also declare that the matter embodied in this project is genuine work done by me and has not been submitted whether to this University or to any other University / Institute for the fulfillment of the requirement of any course of study.

Place : Marwadi University

Date : 09/06/2025

Soumya Bhatt (92320527006) **Signature :** _____

Rabab Travadi (92300527070) **Signature :** _____

Venisha Vora (92300527042) **Signature :** _____

ACKNOWLEDGEMENT

It is indeed a great pleasure to express our thanks and gratitude to all those who helped us. No serious and lasting achievement or success one can ever achieve without the help of friendly guidance and co-operation of so many people involved in the work.

We are very thankful to our guide **Vinod KumarPal**, the person who makes us to follow the right steps during our project work. We express our deep sense of gratitude to for his /her guidance, suggestions and expertise at every stage. A part from that his/her valuable and expertise suggestion during documentation of our report indeed help us a lot.

Thanks to our friend and colleague who have been a source of inspiration and motivation that helped to us during our project work.

We are heartily thankful to the Dean of our department **Dr. R. Sridaran** sir and HoD **Dr. Sunil Bajaja** sir for giving us an opportunity to work over this project and for their end-less and great support to all other people who directly or indirectly supported and help us to fulfil our task.

Soumya Bhatt (92320527006)
Rabab Travadi (92300527070)
Venisha Vora (92300527042)

Signature:_____
Signature:_____
Signature:_____

CONTENTS

Chapters	Particulars	Page No.
1	SYNOPSIS	8
2 2.1 2.2	PREAMBLE General Introduction Module description	9,10
3	REVIEW OF LITERATURE	11
4 4.1 4.2	TECHNICAL DESCRIPTION Hardware Requirement Software Requirement	12
5 5.1 5.2 5.2.1 5.3 5.3.1 5.3.2 5.3.3 5.4 5.4.1 5.4.2	SYSTEM DESIGN AND DEVELOPMENT (Only applicable diagrams) Architectural Design • Class Diagram Dynamic Modeling • Use Case Diagram • Sequence Diagram • Activity Diagram • Any other applicable diagram (applicable) Database Design (If applicable) Relationship Diagram (ER) Menu Design Screen Design	
6	CONCLUSION	13
7	LEARNING DURING PROJECT WORK	14
8 8.1 8.2	BIBLIOGRAPHY Online References Offline References	15

Table Index

Table No.	Title	Page No.
4.1	Hardware Configuration	12
4.2	Software Configuration	12

Figure Index

Table No.	Title	Page No.
5.1	Architectural Design	12
5.2	Class Diagram	12
5.3	Use Case Diagram	12
5.3.1	Sequence Diagram	12
5.3.2	Activity Diagram	12
5.4.1	Relationship Diagram (ER)	12
5.4.2	Menu Design	12
5.4.3	Menu Design (Output)	12
5.4.4	Screen Design	12
5.4.5	Screen Design (Output)	12

Table No.	TITLE	Page No.
5.4.6	Fig- 1	14
5.4.7	Fig- 2	15
5.4.8	Fig-3	16
5.4.9	Fig-4	17
5.4.10	Fig-5	18
5.4.11	Fig-5.1	18
5.4.12	Fig-6	19
5.4.13	Fig-7	20
5.4.14	Fig-8	21
11	Fig-9	22
12	Fig-10	23

1. Synopsis

Cosmetic products such as lipsticks, nail paints, foundations, and eye shadows have become essential parts of modern lifestyles. With the global beauty industry growing rapidly, millions of consumers use makeup products daily without fully knowing the chemicals they are applying to their skin.

While many ingredients are approved for safe use and improve product quality, some chemicals may pose potential health risks, especially with prolonged exposure. Common concerns include allergic reactions, skin irritation, hormone disruption, and in some cases, long-term illnesses.

This project focuses on analyzing the dataset “chemicals_in_makeup.csv” using R programming. The dataset contains details about product names, chemical ingredients, categories (lip, nail, etc.), safety IDs, and counts of chemicals per product. By applying statistical and visual analysis, we aim to:

- Identify the most common chemicals in makeup.
- Detect products/categories with missing or harmful data.
- Understand chemical distribution in various cosmetic categories.
- Visualize patterns using pie charts, bar graphs, histograms, line graphs, and boxplots.

The project helps in building awareness among consumers about harmful trends and encourages safer cosmetic choices. It also demonstrates how data analysis tools like R can support better product evaluation.

2. Preamble

2.1 General Introduction

Data Collection

- **Source:** Public cosmetic safety repository.
- **Dataset File:** `chemicals_in_makeup.csv`.
- **Contents:** Product names, chemical ingredients, categories (nail, lip, etc.), safety ID (CSFID), subcategories, brand details, and chemical counts.
- **Purpose:** To create a foundation for meaningful analysis by using real-world cosmetic data.

Data Preparation

- **Filtering:** Subset and sorting functions were applied to clean the data.
- **Sorting:** Products were alphabetically sorted.
- **Missing Values:** Products with missing CSFID values were identified for careful observation.
- **Purpose:** Ensure the dataset is complete, structured, and ready for analysis.

Data Analysis

- **Techniques Used:**
 - Frequency distribution of products.
 - Identification of missing and duplicate values.
 - Categorization of products by type (lip, nail, foundation, etc.).
- **Purpose:** To uncover meaningful insights and support better visualization.

2.2 Module Description

Visualization:

R programming was used for data visualization. Different graphs such as pie charts, bar charts, histograms, line graphs, and boxplots were prepared. These visuals make complex data easier to understand and interpret.

Interpretation:

The results were studied carefully to highlight:

- Most frequently used chemicals.
- Popular product categories and brands.
- Risk factors due to high chemical counts.
- Distribution of chemicals across different cosmetic product.

3. Review of Literature

Cosmetic safety has been a widely researched area in public health. Many studies suggest that prolonged exposure to certain chemicals in cosmetics can lead to side effects such as skin damage, hormonal imbalance, and allergies.

Research also shows that while regulatory bodies approve ingredients, consumers often lack awareness of harmful effects due to cumulative chemical use. For example, parabens and phthalates are linked to hormone disruption, while heavy metals like lead in lipsticks have raised global concerns.

Online databases and repositories now provide datasets about chemical usage in cosmetics, which researchers and developers can use to analyze safety. Previous works highlight the importance of:

- Developing simple tools for chemical awareness.
- Using data visualization to communicate risks to consumers.
- Encouraging companies to move towards natural and safe ingredients.

This project builds upon such research and provides a practical demonstration of analyzing real-world cosmetic datasets using **R programming**.

4. Technical Description

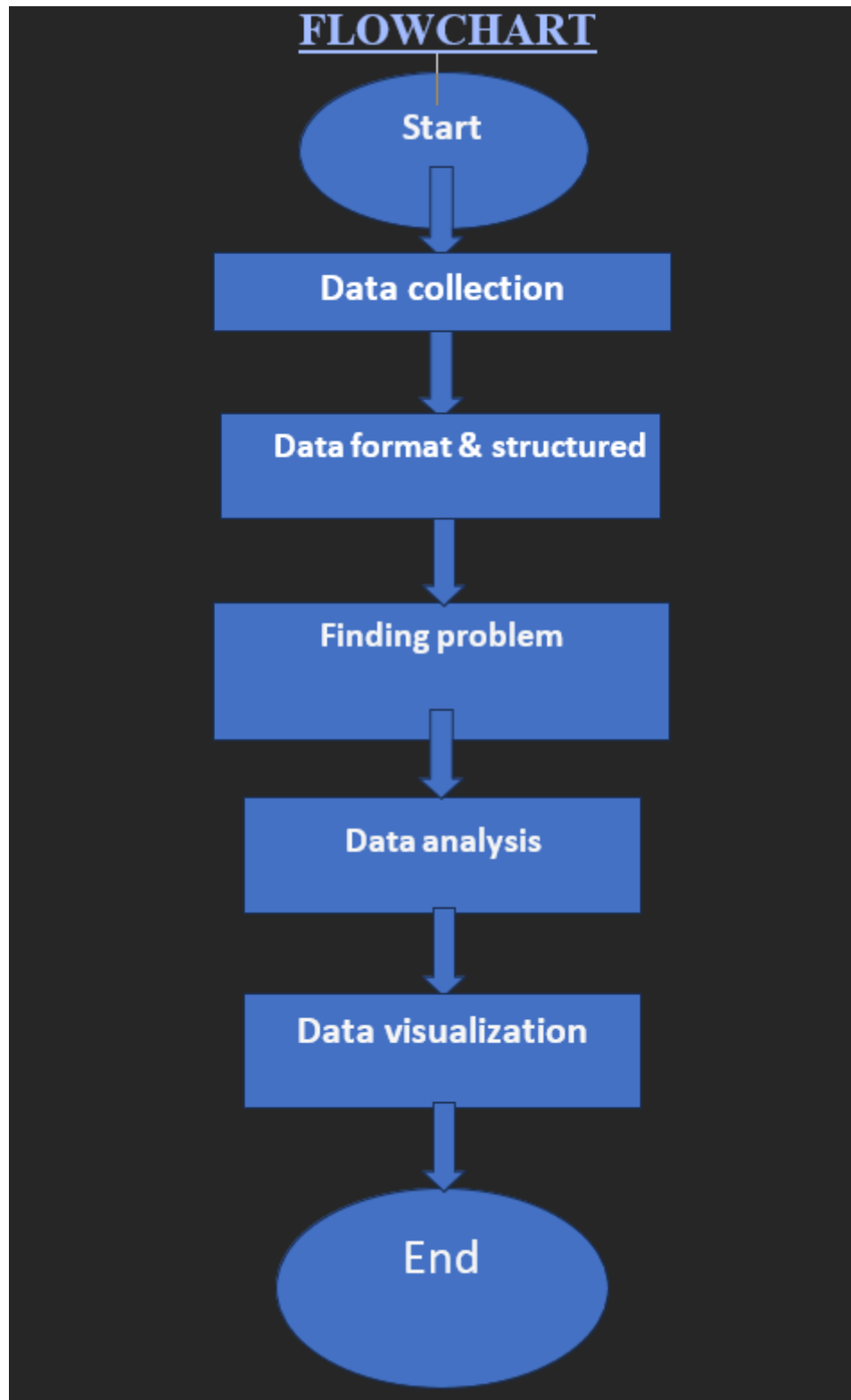
Component	Minimum Specification
Processor	Intel Core i3 (or higher)
RAM	4 GB (or above)
Storage	1 GB free space
Operating System	Windows / Linux / Mac

Table 4.1 Hardware Configuration

Category	Specification
Programming Language	R
Libraries Used	Readr, Plotrix
Dataset	chemicals_in_makeup.csv
Tools	RStudio / Jupyter Notebook

Table 4.2 Software Configuration

5.3.1 Sequence Diagram



5.4.4 Screen Design

Code and its output:

1. Most Common Chemical appearing in the makeup?

Code:

```
print("R1: The chemical appearing the most -")  
  
r1 = max(cc$ChemicalName)  
  
> print("R1: The chemical appearing the most -")  
[1] "R1: The chemical appearing the most -"  
> r1=max(cc$ChemicalName)  
> print(r1)  
[1] "Vitamin A palmitate"
```

Fig-1

2. 1. For easy finding, Sorting based on product name and then finding specific product name:

Code:

```
r2 <- cc[order(cc$ProductName), ]  
  
R2 <- subset(r2, ProductName == "BlueFX")  
  
print("r2: Sorting based on product name -")  
  
print(r2)  
  
print(" R2: Filtered BlueFX entries after sorting:")  
  
print(R2)
```

```

[ reached 'max' / getOption("max.print") -- omitted 114590 rows ]
> print(" R2: Filtered BlueFX entries after sorting:")
[1] " R2: Filtered BlueFX entries after sorting:"
> print(R2)
  CDPHId ProductName CSFId          CSF CompanyId
15    13      BlueFX  317    ULTRAMARINE BLUE      11
16    13      BlueFX  318              MICA      11
17    13      BlueFX  319 FERRIC FERROCYANIDE      11
      CompanyName BrandName PrimaryCategoryId PrimaryCategory
15 OPI PRODUCTS INC. ABSOLUTE FX          59 Nail Products
16 OPI PRODUCTS INC. ABSOLUTE FX          59 Nail Products
17 OPI PRODUCTS INC. ABSOLUTE FX          59 Nail Products
  SubCategoryId          SubCategory CasId CasNumber
15          60 Artificial Nails and Related Products 656 13463-67-7
16          60 Artificial Nails and Related Products 656 13463-67-7
17          60 Artificial Nails and Related Products 656 13463-67-7
  ChemicalId ChemicalName InitialDateReported
15      17918 Titanium dioxide          7/14/2009
16      17918 Titanium dioxide          7/14/2009
17      17918 Titanium dioxide          7/14/2009
  MostRecentDateReported DiscontinuedDate ChemicalCreatedAt
15          12/30/2010              6/9/2010
16          12/30/2010              6/9/2010
17          12/30/2010              6/9/2010
  ChemicalUpdatedAt ChemicalDateRemoved ChemicalCount
15          6/9/2010              1
16          6/9/2010              1
17          6/9/2010              1
> |

```

Fig-2

3. As the Nail art is currently in demand, we have analyzed rows where primary category is nail products.

Code:

```
print("R3: Primary categories makeup (temporary)")
```

```
r3=subset(cc,PrimaryCategory == "Nail Products")
```

```
print(r3)
```

```
> # R3: As the Nail art is currently in demand, we have analyzed
> #rows where primary category is nail products
> print("R3: Primary categories makeup (temporary)")
[1] "R3: Primary categories makeup (temporary)"
> r3=subset(cc,PrimaryCategory == "Nail Products")
> print(r3)
```

CDPHId	ProductName
7	ABSOLUTE Precision Color Powder System - All Shades
8	ABSOLUTE Precision Color Powder System - All Shades
9	ABSOLUTE Precision Color Powder System - All Shades
10	ABSOLUTE Precision Color Powder System Opaque Pink
11	ABSOLUTE Precesion Color Powder System Perfect White
12	ABSOLUTE Precision Color Powder System Translucent Pink
13	ABSOLUTE FX Chromatonex Precision Color Powders
14	ABSOLUTE FX Chromatonex Precision Color Powders
15	BlueFX
16	BlueFX
17	BlueFX
18	GreenFX
19	GreenFX
20	GreenFX
21	GreenFX
22	I'm not Really a Waitress
23	I'm not Really a Waitress
24	I'm not Really a Waitress
25	I'm not Really a Waitress
26	Kyoto Pearl
27	Kyoto Pearl
28	Radiant RedFX
29	Radiant RedFX
30	Radiant RedFX
39	Entity One Nudite Pink UV Gel
40	Entity One UV gels - white, natural white, clear and pink
41	Entity Enamel Base Coat
42	Nudite Nail Sculpting Powders (6)
43	Clear Nail Sculpting Powder
44	Natural Nail Sculpting Powder
45	White Nail Sculpting Powder
46	UV Nail Topcoat
65	Gel Polishes
66	Gel Polishes
67	Gel Polishes
68	Gel Polishes
69	Gel Polishes
70	Gel Polishes
71	Gel Polishes
72	Gel Polishes
73	Gel Polishes
74	Gel Polishes

FIG-3

4. 1. What are the top 5 frequently reported ?

Code:

```
print("R5: Top 5 Most Frequently Reported Products")  
r5 <- sort(table(cc$ProductName), decreasing = TRUE)  
print(head(r5, 5))  
  
print("R5: Top 5 Most Frequently Reported Products")  
1] "R5: Top 5 Most Frequently Reported Products"  
r5 <- sort(table(cc$ProductName), decreasing = TRUE)  
print(head(r5, 5))  
  
Eyecolor      766      Nail Lacquer      667      BLACK LABEL LIPSTICK      487      NYX GIRLS NAIL POLISH      462      ROUND LIPSTICK      427
```

FIG-4

5. Finding products with missing values in CSFID

Code:

```
print("R5: Products with missing values-")  
r5=subset(cc, is.na(CSFId))  
print(r5)
```

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Source
R - R442 - RStudioProject.R
Console Terminal Background Jobs
[1] reached 'max' / getoption("max.print") -- omitted 60118 rows
> # R5: FIND PRODUCTS WITH MISSING VALUES IN CSFID
> print("R5: Products with missing values-")
[1] "R5: Products with missing values-"
> r5=subset(cc, is.na(CSFid))
> print(r5)
  CDPHid      ProductName CSFid CSF CompanyId
1      2  ULTRA COLOR RICH EXTRA PLUMP LIPSTICK-ALL SHADES  NA      4
2      3  Glover's Medicated Shampoo  NA     338
3      3  Glover's Medicated Shampoo  NA     338
4      4  PRECISION GLIMMER EYE LINER-ALL SHADES  NA      4
5      5  AVON BRILLIANT SHINE LIP GLOSS-ALL SHADES  NA      4
6      6  JILLIAN DEMPSEY FOR AVON CELESTIAL EYESHADOW-ALL SHADES  NA      4
11     9  ABSOLUTE Precesion Color Powder System Perfect white  NA     11
31     18  SPECTRA LASH MASCARA  NA      4
32     24  white Premium Lotion Soap  NA    181
37     29  Aloe Up Lil' Kids SPF 45 Aloe Based Sunscreen Lotion  NA     31
38     29  Aloe Up Lil' Kids SPF 45 Aloe Based Sunscreen Lotion  NA     31
39     30  Entity One Nudite Pink UV Gel  NA     33
40     31  Entity One UV gels - white, natural white, clear and pink  NA     33
41     32  Entity Enamel Base Coat  NA     33
42     33  Nudite Nail Sculpting Powders (6)  NA     33
43     34  Clear Nail Sculpting Powder  NA     33
44     35  Natural Nail Sculpting Powder  NA     33
45     36  white Nail Sculpting Powder  NA     33
46     37  UV Nail Topcoat  NA     33
47     38  Pure Blends Natural Lipgloss-Natural  NA     23
48     41  Aloe Up Lil' Kids SPF 30+ Sunscreen Stick  NA     31
49     42  Aloe up SPF 30 Sunscreen Stick  NA     31
50     43  Skin Prep Scrub  NA     43
51     43  Skin Prep Scrub  NA     43
63     45  Enzymatic Sea Mud - professional only  NA     43

```

FIG-5

```

Source
R - R442 - RStudioProject.R
Console Terminal Background Jobs
46     37  UV Nail Topcoat  NA     33
47     38  Pure Blends Natural Lipgloss-Natural  NA     23
48     41  Aloe Up Lil' Kids SPF 30+ Sunscreen Stick  NA     31
49     42  Aloe up SPF 30 Sunscreen Stick  NA     31
50     43  Skin Prep Scrub  NA     43
51     43  Skin Prep Scrub  NA     43
63     45  Enzymatic Sea Mud - professional only  NA     43
64     46  colorstay 12 Hour Eye Shadow Quad- Copper Spice-champagne 99  NA     23
199    48  French Tip  NA     46
200    49  Natural Tip  NA     46
201    50  Perfect white  NA     46
202    51  Swedish white  NA     46
203    52  Pink  NA     46
211    57  colorstay 12 Hour Eye Shadow Quad- peach 98  NA     23
212    58  colorstay 12 Hour Eye Shadow Quad- copper 97  NA     23
213    59  colorstay 12 Hour Eye Shadow Quad- bronze 96  NA     23
214    60  colorstay 12 Hour Eye Shadow Quad- Coffee bean-iced mocha 95  NA     23
215    61  colorstay 12 Hour Eye Shadow Quad- dark brown 97  NA     23
216    62  colorstay 12 Hour Eye Shadow Quad- midtone brown 98  NA     23
217    63  colorstay 12 Hour Eye Shadow Quad- skintone 99  NA     23
218    64  colorstay 12 Hour Eye Shadow Quad- Neutral khakis-taupe vert 93  NA     23
219    65  colorstay 12 Hour Eye Shadow Quad- champ. Shimmer 94  NA     23
220    66  colorstay 12 Hour Eye Shadow Quad- chocolate brown 95  NA     23
221    67  colorstay 12 Hour Eye Shadow Quad- shroom 96  NA     23
222    68  colorstay 12 Hour Eye Shadow Quad- Sandstorm-89 green shimmer  NA     23
223    69  colorstay 12 Hour Eye Shadow Quad- pewter shimmer 90  NA     23
224    70  colorstay 12 Hour Eye Shadow Quad- metal shimmer 91  NA     23
1      1  CompanyName  BrandName PrimaryCategoryId
2      1  New Avon LLC  AVON 44
3      2  J. Strickland & Co.  Glover's 18
4      3  J. Strickland & Co.  Glover's 18
5      4  New Avon LLC  AVON 44

```

FIG-5.1

6. Which is the Minimum products in primary category used by customers?

Code:

```
print("R6: Least product in Primary category-")
```

```
r6 = min(cc$PrimaryCategory)
```

```

print(r6)

> #R6: MINIMUM PRODUCTS IN PRIMARY CATEGORY-
> print("R6: Least product in Primary category-")
[1] "R6: Least product in Primary category-"
> r6 = min(cc$PrimaryCategory)
> print(r6)
[1] "Baby Products"
>

```

FIG-6

7. PIE CHART representing top 5 products in primary category

Code:

```

r7<- table(cc$PrimaryCategory)[1:5]

print(r7)

pie3D(r7,

      labels = paste(round(100 * r7 / sum(r7),1),'%'),

      col = rainbow(length(r7)),

      main = "Pie chart representing top 5 products in primary
category-")

legend("topright",fill = rainbow(length(r7)),

      legend = names(r7),cex = 0.5,)

```

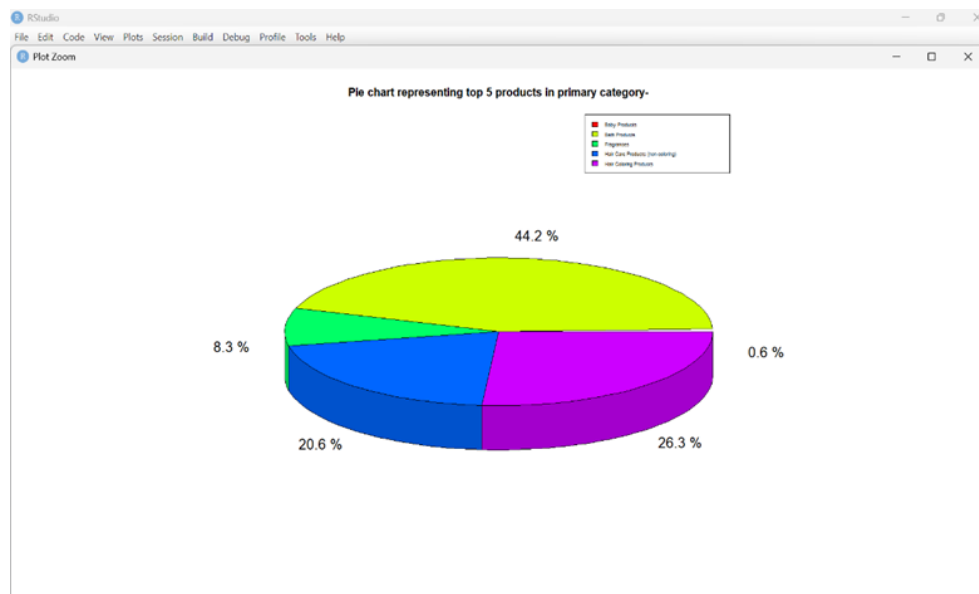


FIG-7

8. BAR GRAPH showing total count of top 7 chemicals.

Code:

```
cc <- read.csv("chemicals_in_makeup.csv")
r8 <- table(cc$ChemicalName)[1:7]
print(r8)
barplot(r8,col = rainbow(length(r8)),
      las = 2,
      main = "Bar graph showing total count of top 7 chemicals-")
```

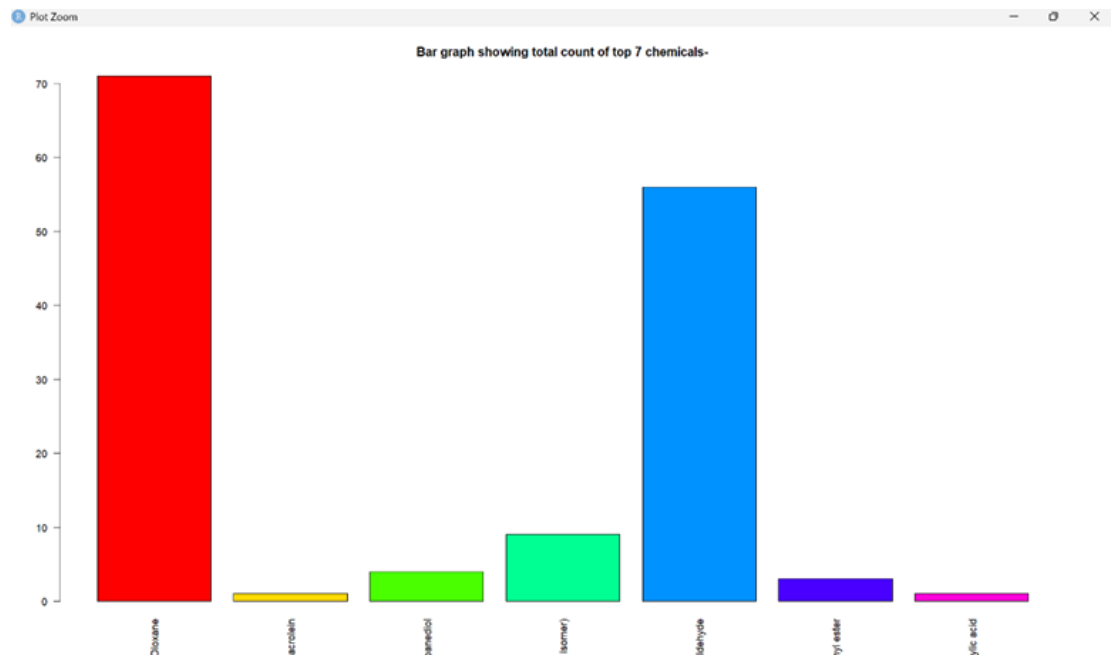


FIG-8

9. HISTOGRAM Chemical Count of Products in subcategory (lip gloss and) chemical MICA

Code:

```
cc <- read.csv("chemicals_in_makeup.csv")

r9 <- subset(cc, SubCategory == "Lip Gloss/Shine" &
ChemicalName == "Mica")

print(r9)

r9$ChemicalCount <- as.numeric(r9$ChemicalCount)

hist(r9$ChemicalCount,
```

```

    main = "Chemical Count of Products in subcategory (lip
gloss) and chemical MICA",

    xlab = "Chemical Count",

    col = "lightblue",

    breaks = 10)

```

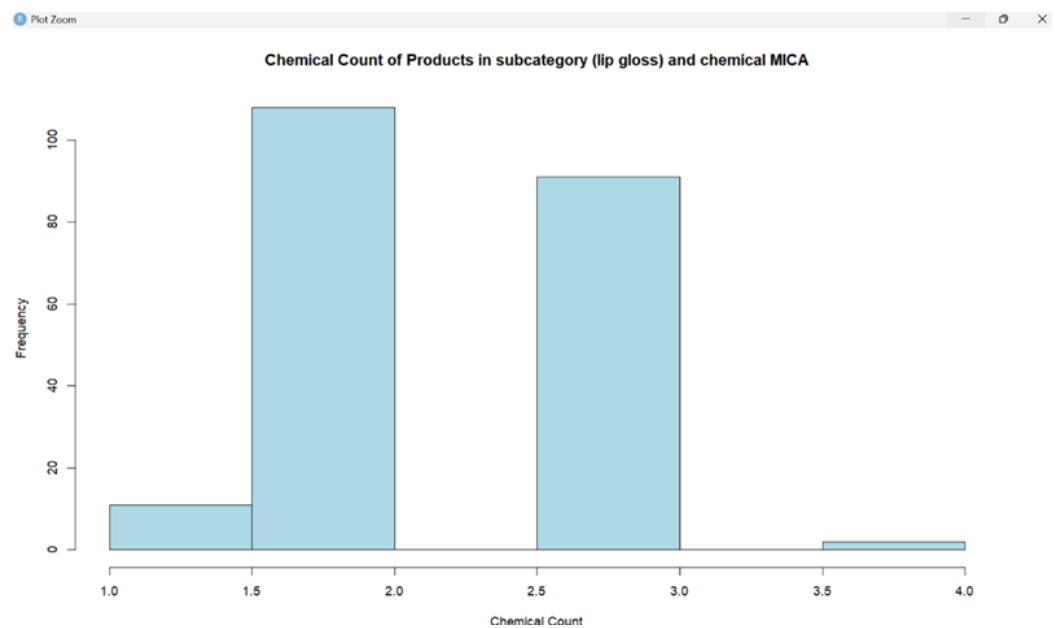


FIG-9

10. LINE GRAPH- Group the data by ProductType and count the products

Code:

```

x <- tapply(cc$CDPHId,cc$ChemicalCount, mean)

print(x)

plot(x,type = 'l',col = 'red',xlim = c(0,10),

     xlab = "Chemical Count",ylab = "Chemical CDPHID",

     main= "Chemical count by CDPHID")

grid()

```

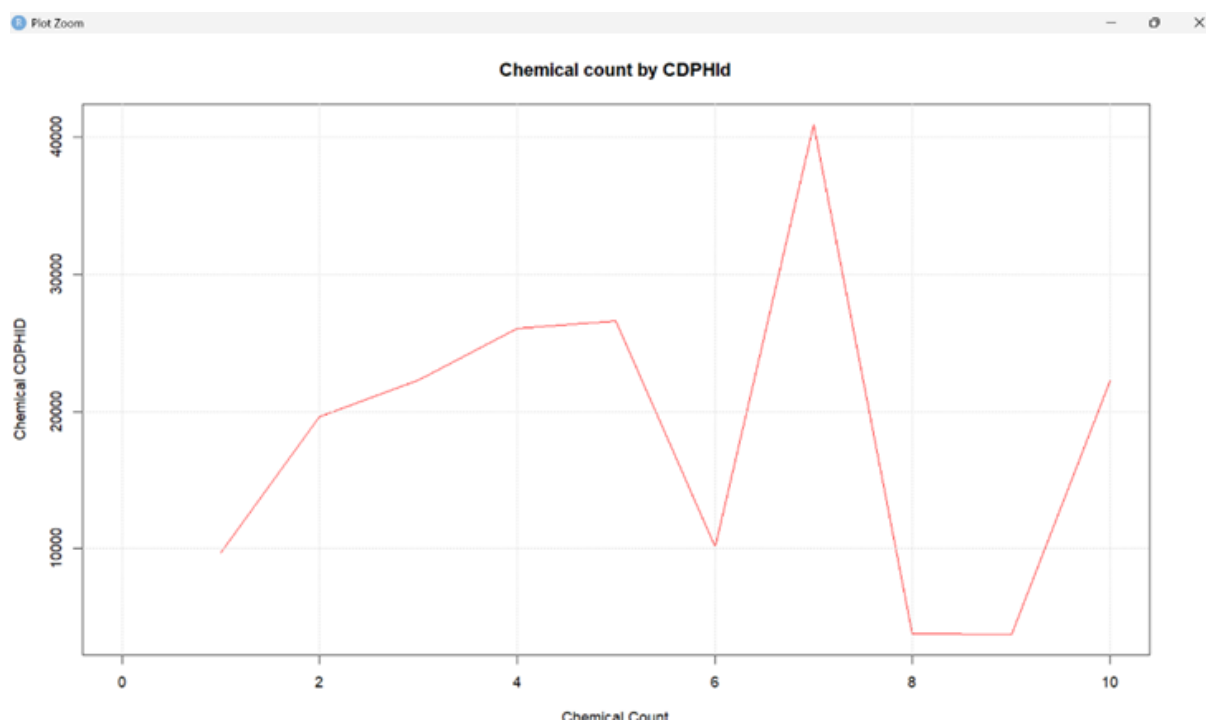


FIG-10

6. CONCLUSION

The project demonstrates how cosmetic safety datasets can be analyzed using R programming. The analysis revealed:

- The most common chemicals in makeup products.
- Nail products and lip products are among the most chemically dense categories.
- Several products had missing chemical safety IDs, which require further investigation.
- Some brands and categories dominate the market, leading to frequent consumer exposure.

- Visualizations such as pie charts, bar graphs, and boxplots helped identify patterns clearly.

Thus, the project contributes towards consumer awareness and highlights the importance of transparency in the cosmetic industry.

7. LEARNING DURING PROJECT WORK

- Gained hands-on knowledge of **R programming** for data analysis.
- Learned how to clean, filter, and sort real-world datasets.
- Practiced creating **different graphs** (pie, bar, histogram, line, boxplot).
- Understood the importance of **data visualization in decision making**.
- Improved teamwork, documentation, and problem-solving skills.

8. BIBLIOGRAPHY

8.1 Offline References

- [1] R. L. P. Adams, *Data Science with R: A Hands-On Guide*. New Delhi: Tech Press, 2023.
- [2] Classroom notes from Prof. Vinod KumarPal, Marwadi University, Rajkot, Gujarat, India, 2025.

8.2 Online References

- [3] CR Cosmetics Safety Repository, *Makeup Chemicals Dataset*. [Online]. Available: Public cosmetic dataset archive.
- [4] R Documentation, *readr package*. [Online]. Available: <https://cran.r-project.org/web/packages/readr/>
- [5] Plotrix Documentation, *Pie3D function*. [Online]. Available: <https://cran.r-project.org/web/packages/plotrix/>
- [6] Research articles on harmful chemicals in cosmetics: <https://www.ncbi.nlm.nih.gov/pmc/>